

December 25, 2016

Searches for $t\bar{t}H$ and tHq with $H \rightarrow \text{leptons}$

CHARLES MUELLER

ON BEHALF OF THE ATLAS AND CMS COLLABORATIONS

Department of Physics and Astronomy

University of Notre Dame, Notre Dame, IN 46556, USA

The latest results from searches for a Standard Model Higgs boson produced in association with a top quark pair ($t\bar{t}H$) and single top quark (tHq) decaying to final states with multiple leptons are presented using datasets from the CMS and ATLAS experiments.

PRESENTED AT

9th International Workshop on Top Quark Physics
Olomouc, Czech Republic, September 19–23, 2016

1 Introduction

While the discoveries of the top quark and Higgs boson were essential in verifying the standard model (SM) of particle physics, many important questions remain unanswered. Specifically, why SM particles have the masses that are observed, and whether or not the top quark’s large mass comes only from its interaction with the Higgs. The top quark is the heaviest fundamental particle, suggesting it could play a special role in electroweak symmetry breaking. According to the SM, the top quark’s large mass comes from its Yukawa interaction with the Higgs. This property of the top quark makes $t\bar{t}H$ and tHq processes an excellent place to test the SM, while remaining sensitive to new physics beyond the SM. By comparing the $t\bar{t}H$ coupling with SM Higgs production via a gluon-gluon loop with virtual top exchange, limits can be set on new physics in the gluon-gluon loop.

The analyses presented here cover recent 2016 LHC Run II $t\bar{t}H$ measurements, as well as previous tHq results by both the CMS [1] and ATLAS [2] collaborations. The Higgs decays targeted include WW^* , ZZ^* , and $\tau\tau$. The Higgs and top systems may decay semi-leptonically or fully leptonically, but the final state leptons must originate from both the top and Higgs systems. The experimental signatures include two same-sign leptons and greater than or equal to three leptons, where leptons are defined as muons or electrons. These requirements narrow the available phase space, which makes signal processes rare, but eliminates large backgrounds.

2 $t\bar{t}H \rightarrow \text{leptons}$

2.1 Object and event selection

The CMS [3] and ATLAS [4] $t\bar{t}H$ analyses begin with a similar object selection strategy. The most important aspect of the object selection is the leptons. The lepton selection uses 3 distinct classes of increasingly tight categories. The first category of leptons is the preselection, consisting of cuts on kinematics, isolation, vertexing and experiment-specific identification variables. The next class is the fakeable selection, which is enriched in non-prompt leptons originating from b-quark decays. These leptons are used later for defining control regions. The final lepton category, known as the “tight” selection is used to select prompt leptons for the signal region definitions. This selection includes tightened versions of the preselection cuts, and additionally, for CMS only, a cut on a multivariate analysis (MVA) value used to identify prompt from non-prompt leptons. This is the primary difference between the ATLAS and CMS tight lepton selections; instead of a cut on a MVA value, ATLAS uses tighter cuts on kinematics and isolation. The remaining objects include jets, taus, and missing transverse energy (MET). These object definitions and selections are similar between CMS and ATLAS.

The event selections are also similar between ATLAS and CMS, and are primarily defined by the tight lepton multiplicity. The selection with the highest statistics is the two-lepton same-sign selection with no hadronically decaying taus. In addition to the nominal requirements, this selection requires lepton transverse momenta be greater than 25 GeV and the presence of at least 4, 5 jets for CMS, ATLAS analyses respectively. The next selection is identical to the previous, but requires exactly 1 hadronic tau. Next are the three-lepton and four-lepton categories. While the ATLAS analysis uses each of these, CMS combines them into a greater than or equal to three-lepton category. Here, the sum of the three lepton electric charges must be ± 1 , and there must be greater than or equal to 2, 3 jets for CMS, ATLAS respectively. Finally, a veto on the presence of two same-flavor opposite-sign leptons within 10 GeV of the Z boson mass is required. An additional requirement for all categories is the presence of b-quark jets. For CMS, this means at least 2 jets passing the loose working point of the b-tagging MVA discriminant, or at least one jet passing the medium working point. For ATLAS, this corresponds to at least one jet passing the medium working point of a b-tagging discriminant.

2.2 Signal extraction and background estimation

While the object and event selections are comparable between ATLAS and CMS, the signal extraction approaches differ significantly. With tighter object definitions and more selective signal regions resulting in a higher signal purity, ATLAS employs a cut and count approach in each of the six categories described in the previous section.

The signal extraction strategy for CMS relies on a two-dimensional MVA technique. There are two boosted decision tree (BDT) MVAs, each trained to discriminate against a single background. The backgrounds targeted individually are the two largest backgrounds, $t\bar{t}V$, where V is a vector boson, and the fake lepton background from $t\bar{t}$. The output from each BDT is plotted on a separate axis, forming a two-dimensional shape, which is then binned, forming a one-dimensional shape in each category. In addition to the categories described previously, the CMS analysis splits the categories described previously by the sign of the sum of the lepton charges, and by the presence (or absence) of two medium b-tagged jets, referred to as the b-tight and b-loose categories. This additional categorization provides additional discrimination power. The BDTs are also trained separately in two-lepton same-sign, and three-lepton categories. The inputs for these BDTs consist of lepton and jet kinematics, solid angles, and MET variables. In addition to these variables, the outputs from a Matrix Element Method (MEM) are also used as inputs to the BDT trained against the $t\bar{t}V$ background (for the three-lepton category only). The MEM output is calculated using a signal hypothesis of $t\bar{t}H, H \rightarrow WW$ assuming three final state leptons. The background hypothesis is $t\bar{t}V$.

In contrast to the signal extraction techniques, the background estimation is simi-

lar between the ATLAS and CMS analyses. The main backgrounds can be classified as reducible and irreducible. The reducible backgrounds consist primarily of non-prompt leptons from b-quark decays in semi-leptonic $t\bar{t}$, but also include prompt leptons with incorrectly measured electric charge. Most importantly, the reducible backgrounds are estimated via data-driven control regions. The irreducible backgrounds are estimated from Monte Carlo (MC) samples and consist of $t\bar{t}V$ and diboson processes. The irreducible backgrounds are referred to as such because the signal regions do not explicitly veto events from these processes.

2.3 Results

Both CMS and ATLAS results were produced with the 2015 and 2016 13 TeV LHC Run II datasets and are summarized below in Figure 1. The ATLAS analysis analyzed 13.2 fb^{-1} of data, while the CMS analyzed 15.2 fb^{-1} of data. CMS measured the best-fit signal strength, $\mu = 2.0^{+0.8}_{-0.7}$, corresponding to a significance of 3.2σ under the background-only hypothesis. ATLAS measured the best-fit $\mu = 2.5^{+1.3}_{-1.1}$, corresponding to a background-only significance of 2.2σ . The leading systematic uncertainties in each analysis include the non-prompt background estimation, jet-vertex association and pile-up modeling, luminosity, signal and background modeling, and the jet energy corrections.

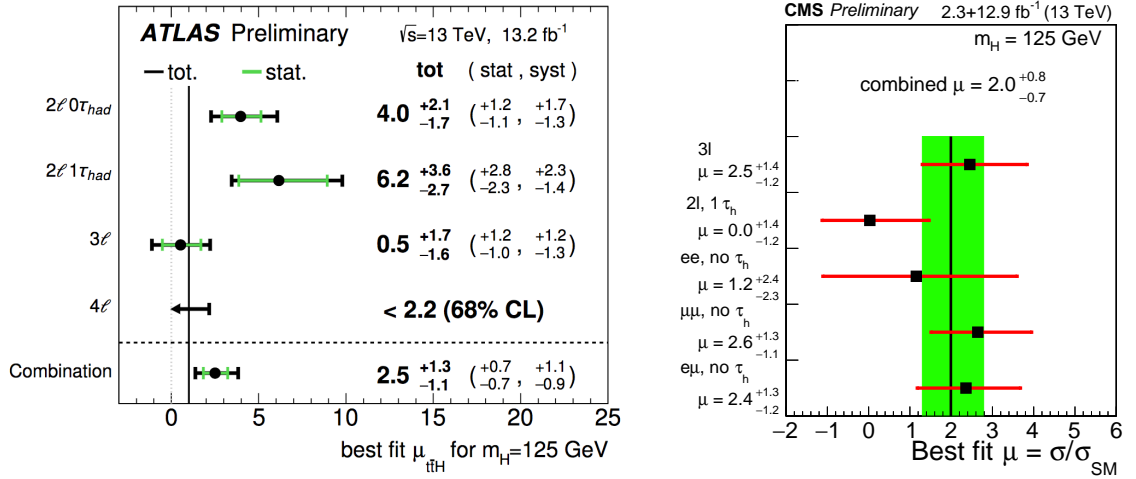


Figure 1: A summary of the latest $t\bar{t}H$ results for ATLAS (left) and CMS (right).

3 $tHq \rightarrow \text{leptons}$

The most recent $tHq \rightarrow \text{leptons}$ measurement was made by CMS at 8 TeV [5]. This search targeted the WW and $\tau\tau$ decays of the Higgs, and measured the interference

between the two leading tHq production diagrams below in Figure 2. This interference is highly sensitive to a non-SM inverted top-Higgs Yukawa coupling, leading to a large enhancement in the tHq cross section under the inverted (negative) coupling scenario.

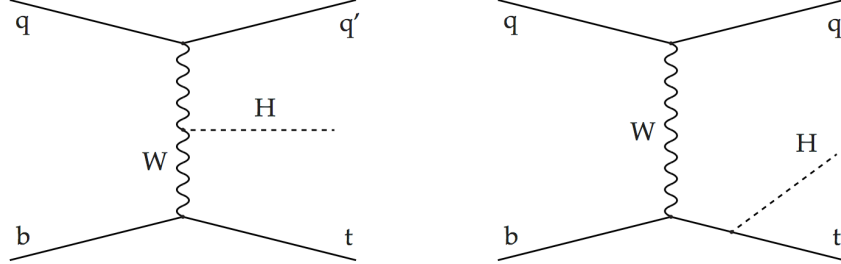


Figure 2: The two interfering tHq process diagrams.

3.1 Object and event selection

The object and event selection use a similar strategy as the $t\bar{t}H$ analysis detailed previously. Prompt leptons are identified with the same MVA mentioned previously, while the other object selections are also very similar to those in the $t\bar{t}H$ analysis.

The event selection consists of a two-lepton same-sign category and a three-lepton category. The two-lepton same-sign category requires lepton transverse momenta greater than 20 GeV, no dilepton mass greater than 20 GeV, no hadronic taus, at least one central jet ($|\eta| < 1.0$), at least one central b-tagged (loose working point) jet, and at least one forward jet ($|\eta| > 1.0$). The three-lepton category requires exactly three leptons with transverse momenta greater than 20,10,10 GeV, the MET must be greater than 30 GeV, and a Z veto is required. In addition to these, all of the jet requirements for the two-lepton same-sign category are required, but the b-tagging working point is tightened from loose to medium. This event selection strategy defines a relatively loose selection, followed by a multivariate signal extraction.

3.2 Signal extraction and background estimation

A linear likelihood discriminator separates signal from background. The signal extraction relies on three types of input variables: forward jet activity, which is only present in the signal process, jet and b-jet multiplicity, and lepton kinematic and charge variables. Two separate likelihoods were used, one for each event selection described previously. The likelihood was trained against loosened selections of $t\bar{t}$, WZ, WWqq, and $t\bar{t}V$ processes for the two-lepton same-sign category. For the three-lepton category, the likelihood was trained against loosened selections of $t\bar{t}$ and $t\bar{t}V$.

The background estimation in the tHq analysis uses the same approach as the $t\bar{t}H$ analysis. The leading backgrounds to tHq are non-prompt (fake) leptons primarily from semi-leptonic $t\bar{t}$ decays, leptons with incorrectly measured charge sign, and $t\bar{t}W$. The non-prompt and charge mis-measured backgrounds are estimated using a data-driven technique via control regions. The remaining backgrounds are estimated from MC.

3.3 Results

The tHq analysis places 95% C.L. upper limits on tHq production under the inverted top-Higgs coupling with 19.5 fb^{-1} of 8 TeV data. The upper limit was measured to be $9.3\text{ (}8.1^{+6.0}_{-11.8}\text{)}$ observed (expected) in the two-lepton same-sign $\mu\mu$ channel, $11.4\text{ (}9.3^{+7.0}_{-13.5}\text{)}$ observed (expected) in the two-lepton same-sign $e\mu$ channel, $11.5\text{ (}8.6^{+6.6}_{-12.4}\text{)}$ observed (expected) in the three-lepton channel, and finally $6.7\text{ (}5.0^{+3.6}_{-7.1}\text{)}$ observed (expected) for all channels combined. The leading systematics include the estimation of the non-prompt background, and the signal and background modeling.

4 Conclusions

The $t\bar{t}H$ and tHq analyses probe the top-Higgs Yukawa coupling at tree-level. The new 2016 $t\bar{t}H$ multilepton results are compatible within approximately 1σ of the SM with $m_H = 125\text{ GeV}$. The ATLAS $t\bar{t}H$ analysis measured a best-fit $\mu = 2.5^{+1.3}_{-1.1}$ corresponding to a significance of 2.2σ . The CMS $t\bar{t}H$ analysis measured a best-fit $\mu = 2.0^{+0.8}_{-0.7}$ corresponding to a significance of 3.2σ . The CMS tHq analysis set 95% C.L. upper limits on the production of the tHq process under the inverted top-Higgs coupling scenario of $6.7\text{ (}5.0^{+3.6}_{-7.1}\text{)}$ observed (expected). These results spell exciting prospects for analyses on the full 2016 dataset and beyond.

References

- [1] CMS Collaboration, JINST 3 (2008) S08004
- [2] ATLAS Collaboration, JINST 3 (2008) S08003
- [3] CMS Collaboration, CMS-PAS-HIG-16-022 (2016)
- [4] ATLAS Collaboration, ATLAS-CONF-2016-058 (2016)
- [5] CMS Collaboration, CMS-PAS-HIG-13-020 (2013)